



UNITED STATES DEPARTMENT OF COMMERCE  
U.S. Census Bureau  
Washington, DC 20233-0001

## 2020 CENSUS PROGRAM MEMORANDUM SERIES: 2022.08

**Date:** March 3, 2022

**MEMORANDUM FOR:** The Record

**From:** Albert E. Fontenot, Jr. **(Signed March 3, 2022)**  
Associate Director, Decennial Census Programs

**Subject:** Census 2020: High-Level Overview of Group Quarters Count Imputation

**Contact:** Andrew Keller  
Decennial Statistical Studies Division  
(301) 763-9308  
[andrew.d.keller@census.gov](mailto:andrew.d.keller@census.gov)

In this memorandum, we provide a high-level discussion of the group quarters (GQ) count imputation process used in the 2020 Census. At the end of data collection, when processing census data from GQs, we realized that many GQs had not provided the necessary information indicating their occupancy status or population count.

To address this issue, we assembled the GQ count imputation (GQCI) team to remove reporting errors from GQs when possible, and to apply a count imputation procedure when valid responses from occupied GQs were not available. This document addresses the steps that the GQCI team took to accomplish these tasks.

The team's work was divided into two stages. First, we partitioned the GQ universe into (a) resolved GQs: those that had a clear status and count, that is, those units classified as nonexistent or vacant GQs with a population count of 0, and occupied GQs with a known positive population count; and (b) unresolved GQs: those that were known to be occupied but did not have a population count. Second, we developed an imputation method that was scientifically sound and could be executed in an efficient manner and applied it to the unresolved GQs.

### 1. Restricting the Imputation Universe

At the beginning of post-data-collection processing, the status of some GQs on Census Day—occupied or vacant—was unclear. As a first step to determine the status of these GQs, an *additional telephone operation* was conducted to contact many of the GQs that provided incomplete data. For some GQs, this effort provided the final status.

The next step to ascertain more GQ statuses encompassed analyzing data provided in previous contacts with the GQ. At the beginning of 2020, mainly in February, the GQ Advanced Contact (GQAC) operation

obtained information that would allow for more efficient enumeration during the GQ Enumeration (GQE) operation around Census Day (U.S. Census Bureau, 2021). Importantly for GQ imputation, the GQAC operation collected the expected population and maximum population capacity expected on Census Day as well as notes from the interviewer regarding the GQ's operational status.

When we initially encountered the problem of GQs that provided a population count of 0 during GQE, we looked at the information obtained during the GQAC operation to determine if the zero count was accurate. For some GQs, the GQAC operation had explicitly indicated that the GQ would have a nonexistent or vacant status on Census Day. If this occurred, we used that GQAC corroborating information to assign a zero population count and called these *GQAC unoccupied resolutions*. A similar issue we faced was a set of GQs in GQAC and GQE for which we completed interviews but did not obtain a population count. Upon careful analysis of the interviewer's notes from the GQAC or GQE or both operations, we inferred that some of these GQs were vacant on Census Day. We assigned these units as *vacant via GQAC and GQE operations*.

In other situations, some GQs had a zero population count because they could not be worked in either GQAC or GQE. As a result, these GQs were determined to be *nonexistent via GQAC and GQE operations*. A critical aspect of this project was understanding when a zero-population count merited no further work. A good example came in the form of GQs that provided services, such as soup kitchens, emergency and transition shelters, regularly scheduled mobile food vans, and targeted non-sheltered outdoor locations. Due to the transient nature of those types of GQs, it was determined that a count of zero obtained during GQE indicated their true status. Consequently, these units were assigned as *vacant due to GQ type*.

An important feature of GQs is that they can be connected as part of larger GQ facilities. Conventional examples of this included universities (many dormitories) or hospitals (several different buildings). Upon inspection, we realized that a few respondents had provided a facility-level count and assigned it to a single GQ in the facility while the other GQs in the facility had no count. Realizing the situation, we allocated the facility-level count across the other GQs in the facility. In short, we *redistributed facility-level counts to unoccupied GQs*. In other facility-level analysis, we discovered duplication of GQs in which one of the duplicates already had a valid population count. As a result, we assigned the other duplicate GQ as *nonexistent due to duplication*. Last, we determined some GQs were *vacant due to expert subject matter review*. For example, some colleges on the list had ceased operations before the 2020 Census—making their GQs vacant.

## **2. Imputing Population for Occupied GQs Without a Size**

Once we identified the unresolved occupied GQs, we imputed a population count. Two aspects shaped the imputation methodology. First, the imputation used prior reported counts from GQAC or other surveys. In particular, the GQAC operation provided expected counts and maximum population capacity counts for Census Day, 2020. We also obtained older GQ counts and maximum counts using response data from previous survey interviews such as the American Community Survey. Examining the set of GQs that had this *auxiliary* GQ information and had reported a valid 2020 count, we observed that the *auxiliary* GQ information was highly predictive of the 2020 GQ count. Second, we used resolved cases of the same GQ type (e.g., colleges, hospitals, prisons) to inform the imputation. For example, if we had to impute a GQ count for a hospital, it was helpful to limit the imputation to use only GQ counts from other hospitals. Last, whenever possible, we restricted the geography as well, using GQ counts only from within the same state.

To perform the imputation, from the GQs that reported a 2020 GQ count, we formed ratios between the auxiliary GQ information and the reported count for each GQ type within the state. Note that, if we had a ratio or auxiliary value that fell outside of a reasonable value, we did not allow it to be used for imputation. To impute a population size for GQs that had some auxiliary counts, we multiplied the auxiliary count by the relevant ratio. For example, if we found that the counts from reporting GQs was, on average, 95 percent of the estimated count from their 2020 GQAC counts for all such hospitals across the state, *for nonresponding GQs* we imputed a 2020 count that was 95 percent of that hospital's estimated count from 2020 GQAC. For the small set of GQs without auxiliary data, other procedures were deployed. In the case of college dorms without such auxiliary data, we used data from the Integrated Postsecondary Education Data System to inform the imputation. Finally, in the absence of all other auxiliary data sources, we imputed the GQ population size using a percentile of the reported population sizes within GQ type and state.

### 3. Conclusion

To determine a GQ operational status or count, two steps were completed. First, we used information obtained from 2020 GQ operations as well as subject matter expertise to identify the operational status. Second, we developed a methodology that relied on auxiliary information about the GQ to impute a final population count for the occupied GQs.

### 4. References

United States Census Bureau (2021). 2020 Census Detailed Operational Plan for: 15: Group Quarters Operation (GQ)," Version 2.0, August 18, 2021.

[2020 Census Detailed Operational Plan: 15. Group Quarters Operation](#) (accessed January 2022).

### The 2020 Census Memorandum Series

The 2020 Census Memorandum Series documents significant decisions, actions, and accomplishments of the 2020 Census Program for the purpose of informing stakeholders, coordinating interdivisional efforts, and documenting important historical changes.

A memorandum generally will be added to this series for any decision or documentation that meets the following criteria:

1. A major program-level decision that will affect the overall design or have significant effect on 2020 Census operations or systems.
2. A major policy decision or change that will affect the overall design or significantly impact 2020 Census operations or systems.
3. A report that documents the research and testing for 2020 Census operations or systems.

Visit 2020 Census on [Census.gov](https://www.census.gov) to access the Memorandum Series, the 2020 Census Operational Plan, and other information about preparations for the 2020 Census.